SELECTING CHILDREN WITH SUSPECTED INFLAMMATORY BOWEL DISEASE FOR ENDOSCOPY WITH THE CALGRANULIN C OR CALPROTECTIN STOOL TEST: PROTOCOL OF THE CACATU STUDY

A. Heida, E. van de Vijver, A.C. Muller Kobold, P.F. van Rheenen

BMJ Open 2017;7:e015636
ABSTRACT

INTRODUCTION The introduction of the faecal calprotectin (FC) test to screen children with chronic gastrointestinal complaints has helped the clinician to decide whether or not to subject the patient to endoscopy. In spite of this, a considerable number of patients without inflammatory bowel disease (IBD) is still scoped. Faecal calgranulin C (S100A12) is a marker of intestinal inflammation that is potentially more specific for IBD than FC, as it is exclusively released by activated granulocytes.

OBJECTIVE To determine whether the specificity of S100A12 is superior to the specificity of FC without sacrificing sensitivity in patients with suspected IBD.

METHODS An international prospective cohort of children with suspected IBD will be screened with the existing FC stool test and the new S100A12 stool test. The reference standard (endoscopy with biopsies) will be applied to patients at high risk of IBD, while a secondary reference (clinical follow-up) will be applied to those at low risk of IBD. The differences in specificity and sensitivity between the two markers will be calculated.

ETHICS AND DISSEMINATION This study is submitted to and approved by the Medical Ethics Review Committee of the University Medical Center Groningen (the Netherlands) and the Antwerp University Hospital (Belgium). The results will be disseminated through a peer-reviewed publication, conference presentation and incorporation in the upcoming National Guideline on Diagnosis and Therapy of IBD in Children.

STRENGTHS AND LIMITATIONS OF THIS STUDY

- Prospective multicentre study evaluating the diagnostic accuracy of a new faecal marker (S100A12) with respect to the currently used faecal marker (calprotectin) to select children with gastrointestinal complaints for endoscopy.
- Our study design reflects current clinical practice in the Netherlands and Belgium.
- Due to the invasive nature of the preferred reference standard (endoscopy) we used clinical follow-up as alternative reference test.
- A limitation of the use of two reference standards is the introduction of a differential verification bias.
- We present a Bayesian approach to deal with the introduced differential verification bias.
INTRODUCTION

Background and rationale
The introduction of the calprotectin stool test to screen children with chronic gastrointestinal complaints has helped the clinician to decide whether or not to refer the patient for endoscopy.\(^1\)-\(^4\) We have shown that children with normal screening test results (\(\leq 50 \, \mu\text{g/g}\)) have a low probability of inflammatory bowel disease (IBD) and should therefore not undergo the invasive reference test (endoscopy) to exclude IBD.\(^5\) Children with elevated calprotectin levels, however, have a high probability of IBD and require referral to an endoscopy unit for endoscopic evaluation of upper and lower gastrointestinal tract.\(^1\),\(^4\),\(^5\) Although use of the calprotectin stool test rarely misses a child with IBD, the number of false positive cases who are scoped is considerable.\(^1\),\(^5\) Calprotectin is a member of the S100 calcium-binding protein family and is a heterodimer of S100A8 and S100A9. The protein is released mainly by neutrophil granulocytes, and also by other activated and damaged cells including monocytes, macrophages and epithelial cells.\(^6\),\(^7\) Calgranulin C (S100A12) is a less investigated member of the S100 protein family.\(^7\),\(^8\) Since S100A12 is only released by activated granulocytes, it is suggested to be more specific for gastrointestinal inflammation caused by IBD than calprotectin.\(^7\),\(^9\)-\(^11\)

Objectives
We hypothesise that a referral strategy based on faecal S100A12 will reduce the number of children wrongly selected for endoscopy as compared with a calprotectin-based strategy. The primary objective is to determine whether the specificity of S100A12 is superior to the specificity of calprotectin without sacrificing sensitivity. The secondary objective is to calculate the diagnostic accuracy characteristics and best cut-offs for both S100A12 and calprotectin.

METHODS

Design
The CACATU (Calprotectin or Calgranulin C test before undergoing endoscopy) study is a prospective, observational, multicentre, diagnostic accuracy study with a paired design. A cohort of children with suspected IBD is screened with the calprotectin stool test (existing test) and with the S100A12 stool test (new test). Confirmation of the target condition (IBD) is based on endoscopy with biopsies (reference standard) or clinical follow-up (secondary reference standard).
**Study design**

Study participants will be recruited from 15 general teaching hospitals and one academic centre in the Netherlands and from one general hospital and two academic centres in Flanders, Belgium. The names of all participating centres can be found in the trial registry (www.clinicaltrials.gov). The principal investigators at the various sites are general paediatricians or paediatric gastroenterologists. Six participating centres (three academic and three general hospitals) have a paediatric endoscopy unit.

**Eligibility criteria**

Patients were eligible if they were between 6 and 17 years old and presented with at least one major criterion or two minor criteria suggestive of IBD (box 1).

<table>
<thead>
<tr>
<th>Major criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Persistent diarrhoea for more than 4 weeks</td>
</tr>
<tr>
<td>• Recurrent abdominal pain with diarrhoea with at least two episodes in the previous 6 months</td>
</tr>
<tr>
<td>• Rectal blood loss</td>
</tr>
<tr>
<td>• Perianal disease (fistula, deep fissure, abscess)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Minor criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Involuntary weight loss</td>
</tr>
<tr>
<td>• First-degree family member with inflammatory bowel disease</td>
</tr>
<tr>
<td>• Anaemia (haemoglobin &lt;2 SD for age and gender)</td>
</tr>
<tr>
<td>• Increased marker of inflammation (erythrocyte sedimentation rate &gt;20 mm/hour or C-reactive protein &gt;10 mg/L)</td>
</tr>
<tr>
<td>• Extraintestinal symptoms (erythema nodosum, arthritis, uveitis, thromboembolism, aphthous ulcers)</td>
</tr>
</tbody>
</table>

*Box 1: Study inclusion criteria (one major criterion or two minor criteria are required to make the patient eligible for participation in the CACATU study)*

**Outcomes**

The primary outcome is the difference in specificity between faecal calprotectin (FC) and S100A12. Secondary endpoints are the difference in sensitivity and the diagnostic accuracy characteristics (sensitivity, specificity, positive predictive value, negative predictive value, area under the curve, best cut-off point) for both markers individually. All diagnostic accuracy characteristics will be calculated with predefined cut-off points.
that have been documented in the medical literature, and with best cut-off points based on receiver operating characteristic (ROC) curves.

**Intervention**

Patients who fulfill the inclusion criteria will be risk stratified (high vs low risk of IBD) according to presenting symptoms, blood tests and stool calprotectin (figure 1). In general, we expect that those participants with increased calprotectin levels (> 50 µg/g) without colon pathogens are likely to be referred to endoscopy (the preferred reference standard) to confirm or exclude IBD. Patients with a normal stool calprotectin test level are likely to have a low probability of IBD and will be followed clinically to determine the final diagnosis (the alternative reference standard), unless there will be other indications to scope them. Paediatricians will be free to use any diagnostic test, such as coeliac disease screening, breath test or ultrasonography (whichever is deemed suitable).

Figure 1: CACATU study flow
Participant study flow

Eligible participants will be invited for participation by the attending paediatrician. Baseline characteristics, date of birth, major and minor criteria (box 1), use of non-steroidal anti-inflammatory drugs and blood tests (haemoglobin, C-reactive protein, erythrocyte-sedimentation rate, serum alanine transaminase and gamma-glutamyltransferase) will be entered on a study website (figure 2, step 1). Participants will be asked to defecate onto a stool collection sheet held above the toilet water and collect one sample with a screw-top container with spoon (step 2). The stool sample is sent to the Department of Laboratory Medicine of the University Medical Center Groningen (UMCG) in a biomaterial envelope. Immediately after arrival, the stool calprotectin level will be measured. The residue will be split with one-half stored at −80°C for S100A12 batch testing at a later stage, and the other half will be used to determine enteric pathogens with a PCR technique (step 3). The PCR analysis will include Shiga toxin-producing Escherichia coli, E. coli O157 gen, Cryptosporidium, Dientamoeba fragilis, Entamoeba histolytica, Giardia lamblia, Salmonella, Shigella/Enteroinvasive E. coli and Campylobacter. Results of calprotectin test and PCR analysis will be uploaded to the website, and will then be made visible to the local clinician (steps 4 and 5). The paediatrician will receive an email notification with an automated advice on the next best move (step 6). However, the choice of the reference standard (endoscopy or clinical follow-up) is up to the paediatrician’s discretion (step 7).

Timeline

The process from faeces collection to completion of the non-invasive diagnostic work-up is supposed to last no longer than 2 weeks. We will exclude samples with a transport time that exceeds 7 days and we will perform a subanalysis with those samples that are received within 4 days. In case of low risk of relapse, the treating paediatricians will receive a reminder for clinical follow-up 6 months after inclusion. The total running time of the study is 30 months, including 6 months to complete the follow-up.

Sample size

The primary outcome of interest is the difference in specificity between the new test (S100A12) and the established test (FC). If the specificity of S100A12 is superior to the specificity of FC without sacrificing sensitivity, we can replace the old test by the new test. McNemar’s test for paired data will be applied to compare specificities between both tests using a 2×2 table exclusively among non-IBD patients (table 1). Study participants with concordant results ((+, +) or (−, −)) do not distinguish between the two
Figure 2. CACATU study flow from first hospital visit to choice of reference test.

STEP 1: The clinician registers the patient on the study website www.cacatustudie.eu

STEP 2: The patient (or parent) collects the stool specimen and sends it to the hospital laboratory

STEP 3: The lab divides the specimen in three portions: calprotectin and PCR are immediately performed; one tube is stored at --80 degrees for calgranulin C testing

STEP 4: The lab sends the results of calprotectin and PCR to the researcher.

STEP 5: The researcher enters the test results on the website

STEP 6: The clinician receives a notification with the results and an automated advice on the next best move.

STEP 7: Paediatrician decides the next best move:
- In case of high probability of IBD: endoscopy*
- In case of low probability of IBD: clinical follow-up

* The ultimate decision to scope is in the hands of the endoscopist

tests. The only information for comparing the sensitivities and specificities comes from those patients with discordant results ((+, −) or (−, +)). Sample size calculation is based on recommendations in Hayen et al.12 Weighed means of specificity of calprotectin were based on a recently published individual patient data meta-analysis.4 We assumed that faecal S100A12 would lead to a 50% relative improvement of specificity (from 70% to 85%). The prevalence of IBD and non-IBD in the CACATU study cohort is expected
to be similar to the prevalence that we found earlier, as the study participants will come from the same region and comparable eligibility criteria will apply. The sample size calculation was done with Power Analysis and Sample Size software (version 11 for Windows). A sample size of 130 subjects with non-IBD achieves 80% power to detect a difference of 0.15 between the two diagnostic tests whose specificities are 0.70 and 0.85. This procedure used a two-sided McNemar test with a significance level of 0.05. The prevalence of non-IBD in the population is 0.64, and the proportion of discordant pairs is 0.23. We aim to include at least 250 participants in order to correct for participants diagnosed with IBD (estimated 36%) and participants that will be lost to follow-up (estimated 25%).

Table 1: Data table of principle of paired design for the fecal markers calprotectin and calgranulin C.

<table>
<thead>
<tr>
<th>Diagnosis: No IBD</th>
<th>Fecal Calprotectin</th>
<th>Fecal Calgranulin C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Concordant (v)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>Discordant (w)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>v + w</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>Discordant (x)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>Concordant (y)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>x + y</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>v + x</td>
</tr>
<tr>
<td></td>
<td></td>
<td>w + y</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N-</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnosis: IBD</th>
<th>Fecal Calprotectin</th>
<th>Fecal Calgranulin C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Positive</td>
<td>Concordant (r)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>Discordant (s)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>r + s</td>
</tr>
<tr>
<td></td>
<td>Positive</td>
<td>Discordant (t)</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>Concordant (u)</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>t + u</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td>r + t</td>
</tr>
<tr>
<td></td>
<td></td>
<td>s + u</td>
</tr>
<tr>
<td></td>
<td></td>
<td>N+</td>
</tr>
</tbody>
</table>

Null hypothesis $H_0$ (specificity): $w = x$  
Alternative hypothesis $H_1$: $w \neq x$

Null hypothesis $H_0$ (sensitivity): $s = t$  
Alternative hypothesis $H_1$: $s \neq t$

**Recruitment**

We asked all participating centres to predict how many eligible patients they could recruit during the enrolment period. Their estimates were based on the list of diagnoses of the previous year, and their estimated totals convinced us that reaching the target sample size is realistic. Retention will be promoted by sending automated reminders to the treating physician to complete the blanks in blood tests, and to reassess patients with initial low probability of IBD after 6 months.

**Index tests**

*Faecal calprotectin*

FC levels will be measured with the fCal ELISA test of BÜHLMANN Laboratories AG
Protocol of CACATU study

(Schönenbuch, Switzerland) according to the manufacturer’s instructions. A level of 50 µg/g is the predefined cut-off value.2,4,5,13

Faecal calgranulin C (S100A12)

S100A12 levels will be measured by one experienced laboratory technician. The maximal duration of storage of the stool sample in our −80°C freezer is 6 months. Analyses will be performed with a sandwich ELISA, trademark Inflamark (CisBio Bioassays, Codolet, France) on a Dynex DS2 Automated ELISA System (Alpha Labs, Easleigh, UK), according to the manufacturers’ instructions. In summary, after the extraction step, 100 µL of prediluted samples will be transferred in duplicate into the corresponding wells coated with anti- S100A12 monoclonal antibody. Incubation time is 30 min, followed by three washing cycles with Tween 20. The next step is adding 100 µL of the second monoclonal antibody, anti-S100A12 coupled to horseradish peroxidase followed by a second incubation period of 30 min and three washing cycles. Next, 100 µL of the substrate, tetramethylbenzidine, is pipetted in all wells. The wells are protected from light and after 10 min, the sulfuric acid stop solution is added. The absorbance will be read at 450 nm. For each duplicate, the mean optical density will be calculated and a calibration curve will be constructed. The curve will be plotted as a cubic regression with DS-matrix software, version 1.23 (Dynex technologies, Chantilly, USA). Purified human S100A12 will be used as calibrator (included in the kit). The predefined cut-off value of S100A12 is 0.75 µg/g, which is based on a reference value study among 120 healthy school-aged children and adolescents.14

Reference tests

Endoscopy

Endoscopy will be the reference standard for patients at high risk of IBD. This procedure will be performed under anaesthesia by an experienced paediatric gastroenterologist. Ideally, both upper and lower gastrointestinal tract will be evaluated according to the revised Porto criteria,15 and biopsies will be taken from every bowel segment. Histopathological examination will be performed by experienced histopathologists. Endoscopists and histopathologists will have access to clinical information and FC test results, but will be blinded to the results of the S100A12 test.

Clinical follow-up

This secondary reference will be applied to patients at low risk of IBD. Six months after study inclusion, the treating paediatrician will receive a notification to enter a second evaluation of major and minor criteria (box 1). Blood tests will only be repeated when
deemed necessary by the treating paediatrician. In addition, a second faeces sample will be collected and sent in for analysis. Patients who remain suspected of having IBD will be referred for further investigations in second instance. At study closure, one of the researchers (AH) will visit the participating centres to cross-check patient records for the definite diagnosis.

**Rationale for choosing reference standard**

Diagnostic endoscopic evaluation of the upper and lower gastrointestinal tract (including intubation of the terminal ileum) in combination with biopsies is the recommended test to diagnose IBD. In children at high risk of IBD with negative endoscopy, small bowel imaging is encouraged. All of these procedures are invasive and require bowel preparation. Endoscopy is mostly performed with the patient under general anaesthesia. Although complications are rare, endoscopy is a burdensome procedure for a child. We found it unethical to expose children at low risk of IBD to endoscopy. Therefore, we decided to perform a secondary reference standard (clinical follow-up) in patients at low risk for IBD and to adjust for its imperfection.

**Blinding**

Laboratory personnel will be blinded to the patient’s history, and to the results of endoscopy and biopsy. Although calprotectin testing is done within 24 hours after arrival of the faeces specimen and the residue is stored at −80°C for S100A12 batch testing at a later stage, sample labeling could theoretically link both faecal tests to one patient. Endoscopists and histopathologists will have access to clinical information and calprotectin test results, but will be blinded to the results of the S100A12 test.

**Confidentiality and data management**

Consecutive patients participating in the study will receive a unique study number. All demographic and medical data will be entered electronically on the study website by the local investigator and stored linked to this study number. Study investigators will receive access to a secured study website. Local investigators are able to consult only data from participants from their own centre. Faeces samples will be marked with a study number label and sent to the Department of Laboratory Medicine at the UMCG. Results of calprotectin test and PCR for enteric pathogens will be uploaded to the website by the coordinating investigator and will be visible to the local clinician. At the end of the study, the data entered on the study website will be cross-checked with the information in the local Electronic Health Databases. Data will be stored during the study period and until 15 years thereafter. When patients and their parents give permission,
residual faeces will be stored for a maximum period of 15 years for future diagnostic research. The researchers AH, EvdV and PvR will have access to the final trial dataset.

**Statistical methods**

Data analysis will be done with SPSS V.22.0 for Windows (SPSS, Chicago, Illinois, USA). Diagnostic accuracy characteristics (sensitivity, specificity, positive predictive value, negative predictive value) will be presented for both markers individually. We will primarily use pre-specified cut-off points of FC and S100A12. In the second instance, we will use the best cut-off points based on the ROC curves for both FC and S100A12. We will use a Bayesian correction method to adjust for differential verification bias in the two reference standards in relation to latent IBD. Based on clinical experience, we defined a prior distribution. We assume that our reference standard endoscopy has 95%–100% sensitivity and 95%–100% specificity to diagnose IBD, and that our secondary reference standard clinical follow-up will have a sensitivity of 80%–100% and a specificity of 60%–80%. Bayes factors will be calculated using JAGS (‘Just Another Gibbs Sampler’), a free program licensed under GNU General Public License.

**Missing values**

In case the index test and reference standard results are missing, the patient will be excluded from further analysis.

**Ethical approval and dissemination policy**

The study will be conducted according to the principles of the Declaration of Helsinki (59th version, October 2008). The Medical Ethics Review Committee (MEC) of the UMCG is of the opinion that this study does not require approval according to the Dutch Medical Research Involving Human Subject Act (WMO). The MEC of the Antwerp University Hospital approved the protocol. The legal guardian(s) of all participants, as well as children aged 12 and above, will need to give informed consent for participation and for storage of material for future research.

In case of important protocol amendments, both the MEC and trial registry will be informed. The results of the trial will be disseminated through a peer-reviewed publication, conference presentation and incorporation in the upcoming National Guideline on Diagnosis and Therapy of IBD in Children.

**Study status**

The first trial participant was included in September 2014. It is anticipated that the trial will end in the spring of 2017.
DISCUSSION

We aim to further improve the accuracy to distinguish patients with a high risk of IBD from those with a low risk of IBD with the ultimate goal to reduce the number of futile endoscopies. We will compare the established faecal marker FC with the relatively unknown faecal marker S100A12. The FC test has excellent sensitivity for IBD (0.92–0.98),²⁻⁴,¹⁸ but its specificity, with point estimates varying between 0.60 and 0.68,²⁻⁴,¹⁸ leaves a considerable proportion of non-IBD patients being exposed to an invasive procedure. Studies with faecal S100A12 showed diagnostic promise under ideal testing conditions in preselected groups of healthy children and children with IBD.⁹,¹¹,¹⁹ We only know of one report that compared FC and S100A12 in children presenting with gastrointestinal complaints.¹¹ The sensitivity and specificity of S100A12 for detection of IBD were both 97%, where FC had a sensitivity of 100% and a specificity of 67%.¹¹

Methodological biases

In this diagnostic accuracy study, the performance of both stool tests will be assessed by verifying the results against endoscopy. Due to the invasive nature of this diagnostic procedure, verification can be performed only in a subset of patients with a high risk of IBD. An alternative reference test (ie, clinical follow-up) will be used in the remainder of the patients. The drawback of this so-called differential verification design is that the second reference test is of lesser quality. Simply adding the results of these two types of reference tests will lead to biased estimates of the overall test accuracy.¹⁶ We plan to correct for this differential verification bias by using a Bayesian approach, as described by De Groot et al.¹⁷ Second, this study is a real-life study, in which the decision to expose a child to endoscopy is based on the combination of presenting symptoms, physical examination and results of blood and stool tests, as is currently recommended by Dutch and international scientific societies. Blinding treating physicians for the FC results was therefore irrational and impractical. Knowledge of the FC level will influence the physicians’ decision to refer a patient for endoscopic evaluation, which gives rise to a work-up bias.²⁰ Furthermore, endoscopists will not be blinded for the level of FC and therefore this may theoretically affect the endoscopists’ assessment of the endoscopy (diagnostic review bias).

Implications for practice

If S100A12 has a better specificity than FC without sacrificing sensitivity, than S100A12 will be the dominant test to select patients for endoscopy. Replacing FC by S100A12 may then reduce the number of non-IBD patients being subjected to endoscopy. This
will be good news for patients (less invasive tests), clinicians (shorter waiting lists for endoscopy) and health insurance companies (reduction of healthcare costs).
REFERENCES


